Chapter 7 Section 2: Comparing Two Means

- **We want to compare…**

- **Each group is considered…**

- **Responses in each group are…**

2-Sample T-test

Same steps for the test of significance:

1.

2.

3.

4.

5.

2 populations with each of their statistics and parameters... (denoted with numbers)

	Pop. 1	Pop. 2
population mean		
population std. dev.		
sample size		
sample mean		
sample std. dev.		

Hypotheses:

- Are comparing…

Ho:

OR

Ha:

Test Statistic:

t =

P-Value:

- **Degrees of freedom:**
  -
  - **(Textbook, p. 549)**

Conclusion:

Almost the same 2 sentences…

-
-

Calculator:

2-Sample T-Interval

Formula:

- Degrees of Freedom:

INTERPRETATION:

Calculator:

<u>ASSUMPTIONS:</u> (for both 2 sample t-test and t-interval)

*

*

Example:

We are trying out a new teaching style that we think will increase test scores. Below is the data for both the treatment and control groups.

	n	$\bar{x}$	s
Treatment	21	81.48	8.12
Control	23	77.85	7.98

Test these hypotheses

Create and interpret a 95% confidence interval

Worksheet: #1 -- 4

Chapter 7 Section 2: Comparing Two Means

- **We want to compare…**

- **Each group is considered…**

- **Responses in each group are…**

2-Sample T-test

Same steps for the test of significance:

1.

2.

3.

4.

5.

2 populations with each of their statistics and parameters... (denoted with numbers)

	Pop. 1	**Pop. 2**
population mean		
population std. dev.		
sample size		
sample mean		
sample std. dev.		

Hypotheses:

- Are comparing…

Ho:

OR

Ha:

Test Statistic:

t =

P-Value:

- **Degrees of freedom:**
    -
    - **(Textbook, p. 549)**

Conclusion:
Almost the same 2 sentences…

-
-

Calculator:

2-Sample T-Interval

Formula:

- Degrees of Freedom:

INTERPRETATION:

Calculator:

<u>ASSUMPTIONS:</u> (for both 2 sample t-test and t-interval)

*

*

Example:

We are trying out a new teaching style that we think will increase test scores. Below is the data for both the treatment and control groups.

	n	$\bar{x}$	s
Treatment	21	81.48	8.12
Control	23	77.85	7.98

Test these hypotheses

Create and interpret a 95% confidence interval

Worksheet: #1 thru 4

1)

STATE	CHECK
1) 2 indep. SRS	1) assumed
2) 2 normal pop or $n1$ & $n2 \geq 30$	2) n1 and n2 = 30

df = 57.9513

$$(\bar{X}1 - \bar{X}2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (2.916, 6.484)$$

We are 95% confident that the difference between the averages of the number of accidents per month occurring the in the two departments is between 2.916 and 6.484 accidents.

Since 0 (meaning no difference between the two) is not in the interval, we have evidence to say that there IS a difference between the two departments.

2)

STATE
1) 2 indep. SRS
2) 2 normal pop
or
$n1$ & $n2 \geq 30$

CHECK
1) assumed
2) $n1 = 45 > 30$
$n2 = 35$

df= 68.8650

$$(\bar{X}1 - \bar{X}2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (2.4386, 3.3214)$$

We are 90% confident that the difference between the average cost of groceries in suburban and inner city stores is between \$2.4386 and \$3.3214.

Since 0 is not in the interval, we have evidence to say that there is a difference between the suburban and inner city stores

3)

STATE	CHECK
1) 2 indep SRS	1) assumed
2) 2 normal pop	2) n1 = 50 > 30
or	n2 = 40
n1 & n2 $\geq$ 30	

Ho: $\mu_1 = \mu_2$

Ha: $\mu_1 > \mu_2$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = 8.2459$$

P(t > 8.2459| df=77.1146) = 0

- We reject Ho b/c p-value < alpha = 0.05.
- We have sufficient evidence that the average nicotine content of cigarette 1 is greater than the average nicotine content of cigarette 2.

4)

STATE	CHECK
1) 2 indep SRS	1) assumed
2) 2 normal pop or $n1$ & $n2 \geq 30$	2) $n1 = 40$, $n2 = 50$ > 30

$\bar{X}_1 =$

$\bar{X}_2 =$

$S_1 =$

$S_2 =$

$n_1 =$

$n_2 =$

Ho: $\mu_1 = \mu_2$

Ha: $\mu_1 \neq \mu_2$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = 2.6040$$

$2 * P(t > 2.6040 | df = 86.6260) = 0.0108$

- We reject Ho b/c p-value < alpha = 0.05.
- We have sufficient evidence that the average number of business lunches for insurance executives is not equal to the average number of business lunches for banking executives.

POOLED 2 Sample t-test

to pool resources

Notes:

- **2 samples…**

  taken from 2 diff. pop.

- ✳ **Both populations…**

  have same σ

- **How do we estimate** ~~??~~ σ

  $\sigma_1 = \sigma_2$

  $S_1 \quad S_2$

  - Pool both S_1 + S_2
  - combine into 1 estimate for σ.

- **In doing this, we want to…** give more weight to larger n.
- **sp =** on form. sheet
  calculator
- **sp is called the…** pooled std. dev.
  pooled estimator of σ

Statistics:

Population 1

$\bar{x}_1$

s_1 → s_p ← s_2

n_1

Population 2

$\bar{x}_2$

s_2

n_2

Hypotheses:

Same

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \quad \mu_2$

Test Statistic:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$

df = $n_1 + n_2 - 2$

on calc

P-Value:

Same

Conclusion:

- Same
-

Confidence Interval Formula:

$$(\bar{X}_1 - \bar{X}_2) \pm t^* S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Interpretation: Same

On Calculator:

- Same — Pooled: YES
-

Example:

Two companies, A and B, make two different types of cookies. We want to test to see if the mean grams of sugar in company A's cookies are greater than that of Company B. We take 2 independent SRS and find the statistics below: It's believed that the pop. std. dev. are same.

	n	$\bar{x}$	s
A	15	6.21	1.34
B	21	6.02	1.32

Test these hypotheses at the 0.05 significance level.

* Assump

3) $\sigma_1 = \sigma_2$ 3) stated

$H_0: \mu_A = \mu_B$

$H_a: \mu_A > \mu_B$

$$t = \frac{\bar{x}_A - \bar{x}_B}{S_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} = 0.4231$$

→ $P(t > 0.4231 \mid df = 34) = 0.3374$

- fail to reject H_0
- $\mu_A = \mu_B$

Computer output:
turn to p.562 #65

- notice what the output gives you

- ALWAYS gives 2-sided test (specifically, 2-sided p-value)

- how do you get the p-value for a 1-sided test?

Some vocab.... STANDARD ERROR

- the bottom of the test statistic (or the std. dev. of the stat.)